УДК: 004.93 https://doi.org/10.15407/jai2023.03.154

О.С. Стрюк[1], Ю.П. Кондратенко[2]
[1,2]Чорноморський національний університет ім. Петра Могили, Україна
вул. 68 Десантників, 10, м.Миколаїв, 54000
[2]Інститут проблем штучного інтелекту Міністерства освіти і науки України
і Національної академії наук України, Україна
вул. Мала Житомирська, буд. 11/5, м.Київ, 01001
[1]oleksandr.striuk@gmail.com
[2]y_kondrat2002@yahoo.com
[1]https://orcid.org/0000-0002-6391-4382
[2]https://orcid.org/0000-0001-7736-883X

МЕТОДИ ПРИКЛАДНОГО ЗАСТОСУВАННЯ ГЕНЕРАТИВНИХ ЗМАГАЛЬНИХ МЕРЕЖ ПРИ ОБРОБЦІ ГРАФІЧНИХ ДАНИХ

O. Striuk[1], Y. Kondratenko[2]
[1,2]Petro Mohyla Black Sea National University, Ukraine
68-Desantnykiv Str., 10, Mykolaiv, 54003
[2]Institute of Artificial Intelligence Problems of the Ministry of Education and Science of Ukraine
and the National Academy of Sciences of Ukraine, Ukraine
Mala Zhytomyrs’ka Str., 11/5, Kyiv, 01001
[1]oleksandr.striuk@gmail.com
[2]y_kondrat2002@yahoo.com
[1]https://orcid.org/0000-0002-6391-4382
[2]https://orcid.org/0000-0001-7736-883X

METHODS OF APPLIED UTILIZATION OF GENERATIVE ADVERSARIAL NETWORKS IN GRAPHIC DATA PROCESSING

Анотація. У доповіді досліджується важлива область штучного інтелекту — генеративні змагальні мережі (ГЗМ), які використовуються для створення високоякісних штучних зразків даних. ГЗМ зазнали значного розвитку та застосування в різних секторах, включаючи обробку графічних даних. В доповіді зосереджено увагу на прикладному використанні ГЗМ та їхній архітектурі. Розглянуті базові принципи функціонування ГЗМ, висвітлюються переваги та недоліки, включаючи проблеми з навчанням, зникаючі градієнти та конвергентні осциляції, і описуються заходи для подолання цих проблем. Також розглядаються сучасні дослідження в галузі ГЗМ та їхнє застосування у різних галузях, включаючи кібербезпеку, медицину, криміналістику та комп'ютерний зір. Висвітлено практичні результати авторів доповіді щодо власних експериментів з ГЗМ, їх оптимізації та вдосконалення архітектури. Мета дослідження полягає в аналізі архітектурних особливостей ГЗМ з метою покращення процесу їх навчання.

Ключові слова: штучний інтелект, глибоке навчання, нейронні мережі, генеративні змагальні мережі, обробка зображень.

Annotation. The paper explores an important area of artificial intelligence — Generative Adversarial Networks (GANs), which are used to create high-quality artificial data samples. GANs have undergone significant development and application in various sectors, including the processing of graphical data. The report focuses on the practical use of GANs and their architecture. It discusses the fundamental principles of GAN operation, highlights the advantages and disadvantages, including issues with training, vanishing gradients, and convergence oscillations, and describes measures to overcome these problems. It also examines current research in the field of GANs and their applications in various domains, including cybersecurity, medicine, forensics, and computer vision. Practical results from the report's authors regarding their own GAN experiments, optimization, and architecture improvements are presented. The research aims to analyze the architectural features of GANs to enhance their training process.

Keywords: artificial intelligence, deep learning, neural networks, generative adversarial networks, image processing.

Вступ

Генеративні змагальні мережі (ГЗМ) [1] стали трансформаційною силою в царині генеративного штучного інтелекту, змінивши підхід до різних областей, від комп’ютерного бачення до обробки природної мови [2, 3]. Їх здатність створювати штучні зразки, які імітують

реальний розподіл даних, вплинула на інновації та широке застосування в різних секторах. Однією з ключових сфер застосування ГЗМ сьогодні є обробка графічних даних [2].

Спектр застосування ГЗМ має широкий діапазон, що включає генерацію фотореалістичних зображень, створення нових хімічних молекул і матеріалів, покращення інструментальних телескопічних знімків, синтез біометричних даних, збільшення роздільної здатності фотографій, створення і виявлення шкідливого програмного забезпечення, а також виявлення аномалій в різноманітних прикладних галузях (кібербезпека, медицина, зміни клімату, природні науки тощо) [3, 4].

Фокус даної доповіді сконцентровано на обробці графічних даних, зосереджуючись саме на прикладному аспекті використання ГЗМ.

Базис ГЗМ

ГЗМ — архітектура глибокого навчання, яка характеризуються змагальним підходом до навчання та здатністю до генералізації можливостей через трансферне навчання [1].

ГЗМ здатні генерувати високоякісні штучні зразки, які можуть бути використані як додаткові навчальні дані для інших систем штучного інтелекту, наприклад, нейронних мереж, що вирішують завдання бінарної класифікації, сегментації або прогнозування. Додаткові навчальні дані здатні підвищити точність існуючих моделей.

ГЗМ функціонують шляхом навчання двох нейронних мереж: генератора та дискримінатора. Генератор приймає вектор випадкового шуму як вхідні дані та продукує синтетичну вибірку, яка нагадує реальні дані [1].

Дискримінатор приймає як реальні, так і синтетичні зразки як вхідні дані та ідентифікує їх (підробка чи реальні дані). Дві мережі «грають» у змагальний навчальний процес, схожий на гру, коли генератор намагається змусити дискримінатор «повірити», що його синтетичні зразки є «справжніми», а дискримінатор намагається точно ідентифікувати справжні та підроблені зразки (рисунок 1).

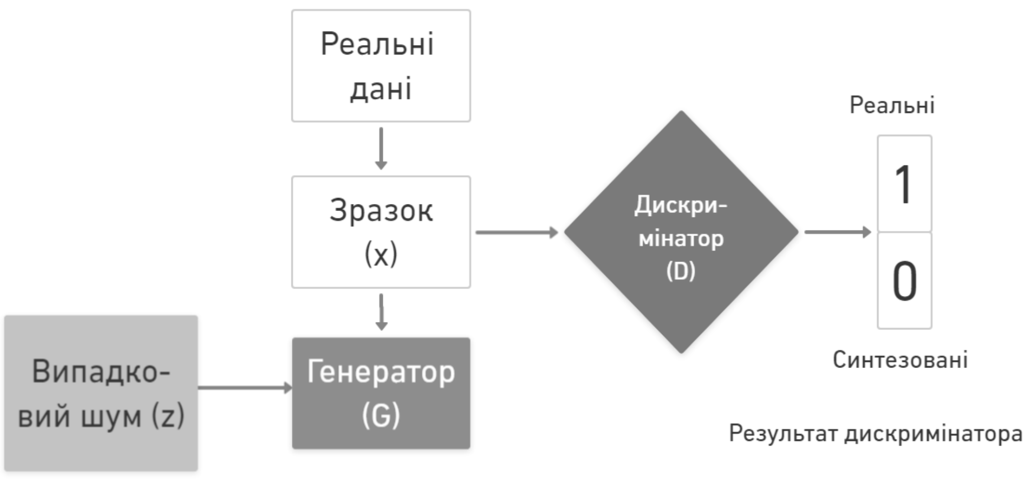


Рис. 1. Структура ГЗМ

Переваги і недоліки ГЗМ

Суттєвою перевагою ГЗМ є те, що вони не вимагають явного моделювання розподілу ймовірностей реальних даних. Це робить цей тип нейронних мереж придатним для генерації багатовимірних, складних даних, таких як зображення та фотографії [5].

Незважаючи на загальну ефективність, процес навчання ГЗМ може бути важкою задачею і вимагати ретельного тонкого налаштування гіперпараметрів моделі [6].

Крім того, ГЗМ можуть страждати

від колапсу режиму, коли генератор виробляє лише обмежений набір зразків, які не в змозі охопити різноманітність спектру реальних даних [7].

Ще однією проблемою є зникаючі градієнти, що можуть виникати, коли дискримінатор стає занадто ефективним у розрізненні справжніх зразків від підроблених, що призводить до труднощів для навчання генератора [5].

Певні конвергентні осциляції можуть також виникати, коли генератор і дискримінатор застрягають у контурі зворотного зв'язку і не можуть знайти оптимальне рішення.

Оцінка показників продуктивності ГЗМ також може бути трудомісткою, оскільки немає абсолютно чіткої метрики для вимірювання якості згенерованих зразків [8].

Для вирішення цих проблем зараз застосовуються різні методи, включаючи використання різних функцій втрат, методів оптимізації, регуляризації та варіативність архітектур [6]. Нижче будуть наведені приклади власних результатів авторів доповіді щодо практичної імплементації цих методів.

Сучасні дослідження в галузі ГЗМ

Окрім загальновідомих сфер застосування ГЗМ, таких як наука, інтерактивні медіа, реклама, відеоігри, сьогодні даний тип штучних нейронних мереж знаходить широке застосування і в інших практичних царинах, як от: кібербезпека, аудіосинтез, трансферне навчання, медицина, криміналістика, 3D-моделювання, виробництво нових ліків та інше [3, 9].

ГЗМ також відіграють важливу роль у дослідженнях у галузі комп'ютерного зору. Вони можуть суттєво поліпшувати зображення, отримані телескопами та розвідувальними дронами, а також використовувати методи для виправлення розмитості, заповнення пробілів та підвищення роздільної здатності. Це призводить до значного покращення результатів існуючих систем глибинного навчання, розширюючи їхні можливості та підвищуючи точність обробки зображень [3].

Усе це свідчить про значущий вплив глибокого навчання та ГЗМ на різні галузі та наголошує їхню роль у розвитку сучасного технологічного світу.

Метою дослідження є аналіз архітектурних особливостей ГЗМ з метою вдосконалення та стабілізації їх процесу навчання, що на сьогодні є найбільшою проблемою цього типу штучних нейронних мереж.

Експериментальні результати ГЗМ

З метою оптимізації стратегії проєктування ГЗМ, яка б забезпечила більш стабільне і ефективне навчання з метою покращення точності моделей, авторами було проведено ряд експериментів по вдосконаленню архітектури ГЗМ.

Конвенційна ГЗМ (англ. Vanilla GAN) [10]. Перша спроєктована модель була створена з метою синтезу реалістичних зображень, схожих на зразки з набору даних MNIST [11], який складається з рукописних цифр. Щоб досягти цього, автори імплементували кілька методів і технік оптимізації. Застосували функції активації LeakyReLU, бінарну крос-ентропійну функцію втрати та одностороннє згладжування міток для стабілізації навчання. Використання згладжування міток передбачає призначення цільового значення 0,9 реальним зображенням замість радикального 1, що може зробити навчання дискримінатора більш надійним. Ми також використали оптимізатор Адам з низькою швидкістю навчання та включили шум як вхідні дані для генератора для поступового генерування зображень. Крім того, ми реалізували функцію візуалізації для моніторингу прогресу навчання та якості згенерованих зразків. Таким чином було проведено як параметричну, так і структурну оптимізацію.

Ми мали на меті навчити модель створювати реалістичні зображення, схожі на рукописні зразки MNIST. Наш підхід передбачав вибір архітектури, нормалізацію, функції активації, функції втрат, згладжування міток і методи оптимізації для підвищення стабільності

навчання та якості зображення, що в кінцевому підсумку спрацювало на досягнення головної мети — створення високоякісних синтетичних зображень [10].

Результати експерименту з конвенційною архітектурою ГЗМ представлені на рисунку 1.



Рис. 2. Зразки, згенеровані ГЗМ після навчання на наборі даних MNIST

Глибока згорткова ГЗМ (англ. DCGAN) для синтезу біометричних зразків відбитків пальців [12]. Друга модель була спроєктована з метою синтетичного відтворення фотореалістичних зображень відбитків пальців, які б за спектром властивостей відповідали реальним біометричним даним.

У рамках даного експерименту були використані різні техніки оптимізації для досягнення найкращого результату в генерації зображень.

Ми спроєктували архітектуру з використанням згорткових нейронних мереж. Генератор був створений з послідовності транспонованих згорткових шарів з функцією активації ReLU та Batch Normalization для стабільності. Дискримінатор також мав згорткові шари з функцією активації LeakyReLU та Batch Normalization.

Також було застосовано бінарну крос-ентропійну функцію втрати. Крім того, були використані різні швидкості навчання окремо для мереж генератора і дискримінатора для оптимізації навчання.

Процес включав у себе аналіз та моніторинг навчання, включаючи відстеження втрат та періодичне збереження згенерованих зображень для візуалізації. Також було використано ініціалізацію ваг для покращення стабільності навчання.

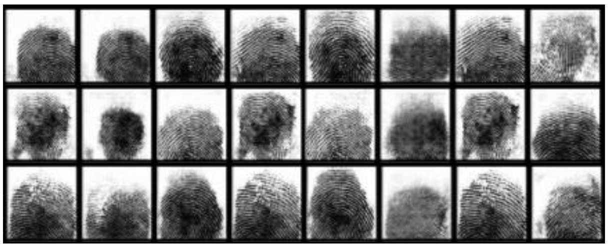


Рис. 3. Згенеровані результати DCGAN після 1000-го навчального циклу

Важливість методу полягає у тому, що він частково розв'язує проблему обмежень використання реальних відбитків пальців для навчання глибоких нейронних мереж. Реальна біометрична інформація належить до особистих даних і може мати обмеження у використанні, що пов’язано з конфіденційною природою цих даних. Штучно створені зображення відбитків пальців не мають таких обмежень і можуть бути використані в дослідженнях та прикладних аспектах, таких як біологічні дослідження, судова експертиза, технологічні методи біометричної безпеки [13]. Розроблена глибока згорткова модель продемонструвала потенціал щодо здатності генерувати реалістичні відбитки пальців.

ГЗМ з покращеною здатністю для збільшення роздільної здатності (англ. ESRGAN) [14]. При проєктуванні власної мобільної модифікації ГЗМ для збільшення роздільної здатності автори використали архітектуру ESRGAN із попередньо навченими ваговими коефіцієнтами та змінними параметрами інтерполяції, що забезпечило плавне керування та точніше налаштування моделі нейронної мережі.

Як Benchmark автори використали два різні набори даних у вільному доступі:

а) Sokoto Coventry Fingerprint Dataset

(SOCOFing) — біометрична база даних відбитків пальців, яка містить 6000 дактилограм; цей набір даних також підходить для інших завдань, пов'язаних із глибинним навчанням [15];

б) набір даних BIRDS 400, розроблений спеціально для класифікації зображень видів птахів [16].

Автори підготували зображення низької роздільної здатності для подальшої обробки на моделі та збільшення роздільної здатності оригінальних зразків у чотири рази без втрати якості.

Розмір усіх тестових зразків було змінено до 96Х96 пікселів (вибране випадковим чином значення) і перетворено у формат PNG за допомогою бібліотеки Python — PIL. Задачею було збільшити роздільну здатність цих експериментальних зразків у 4 рази, тому роздільна здатність кінцевого зображення мала бути 384Х384 пікселя. Після обробки обраних наборів зображень за моделлю автори отримали очікувані результати.

Спочатку було застосовано модель до зразків відбитків пальців. Ось як зразки виглядали до покращення в 96Х96 пікселів, рис. 4.

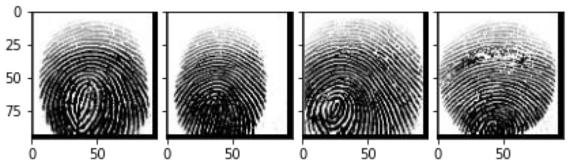


Рис. 4. Зразки SOCOFing перед обробкою, 96Х96 пікселів

А ось як виглядають зразки після збільшення моделлю в 4 рази, до 384Х384, рис. 5.

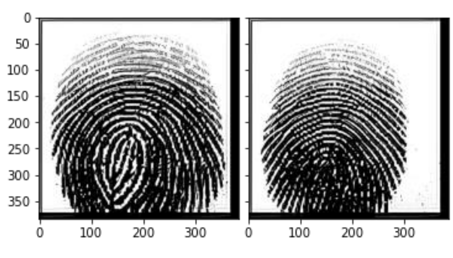


Рис. 5. Зразки SOCOFing після обробки

Нижче наведені результати обробки зразків зображень із набору даних BIRDS 400, рис. 6 і рис. 7 відповідно. Як ми бачимо, зразки були успішно збільшені до 384Х384 пікселів із збереженням максимальної якості зображення.



Рис. 6. Зразок набору даних BIRDS 400 перед обробкою, 96Х96 пікселів



Рис. 7. Зразки набору даних BIRDS 400 після обробки, 384Х384 та 96Х96 пікселів

Експеримент продемонстрував, що модель на основі ГЗМ з коректно налаштованою архітектурою та функцією втрат здатна значно збільшити роздільну здатність обробленого зображення, зберігаючи при цьому якість і стабільну графічну структуру [14].

ГЗМ для виявлення аномалій [17]. Автори статті також розробили концепцію застосування ГЗМ для виявлення аномалій у графічних даних із застосуванням нечіткої логіки (нечітких множин) і лінгвістичних термів для більш ефективної фіксації атипових розподілів даних.

Для експериментального дослідження (proof of concept) також було використано набір даних MNIST. Була розроблена модель на базі конвенційної ГЗМ з використанням середньої абсолютної похибки як інструменту виявлення аномалій, а також з додаванням нечіткого класифікатора виявлених аномалій по лінгвістичним термам — низький, середній, високий, для більш ефективного візуального представлення результатів. Збудована система продемонструвала здатність розрізняти аномалії в графічних даних (рисунки 8 та 9).

В якості «нормальних» даних були обрані усі зразки, що містили цифру «1», а всі інші цифри були визначені як аномальні. Модель вірно класифікувала

більшість зразків із точністю 95%.

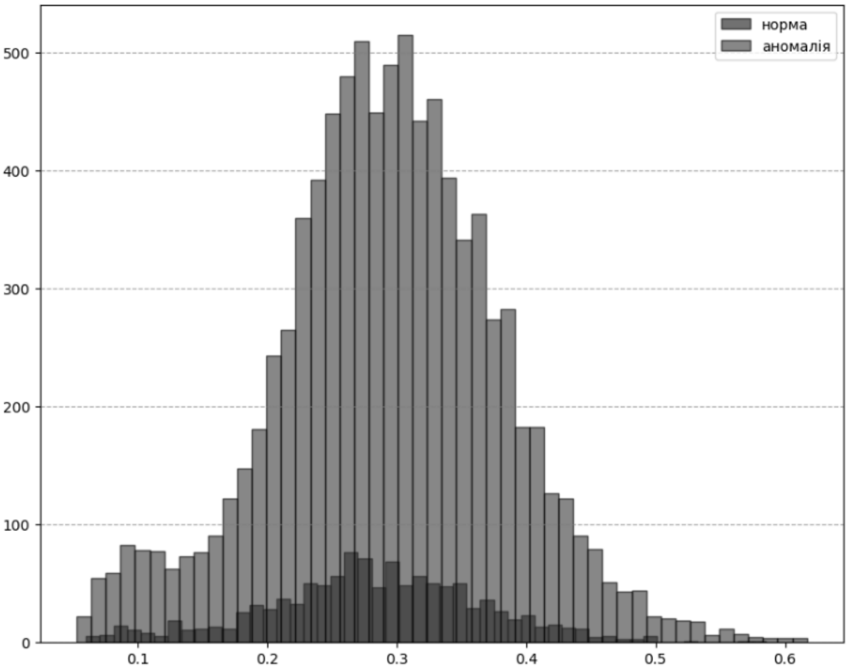


Рис. 8. Розподіл балів аномалії

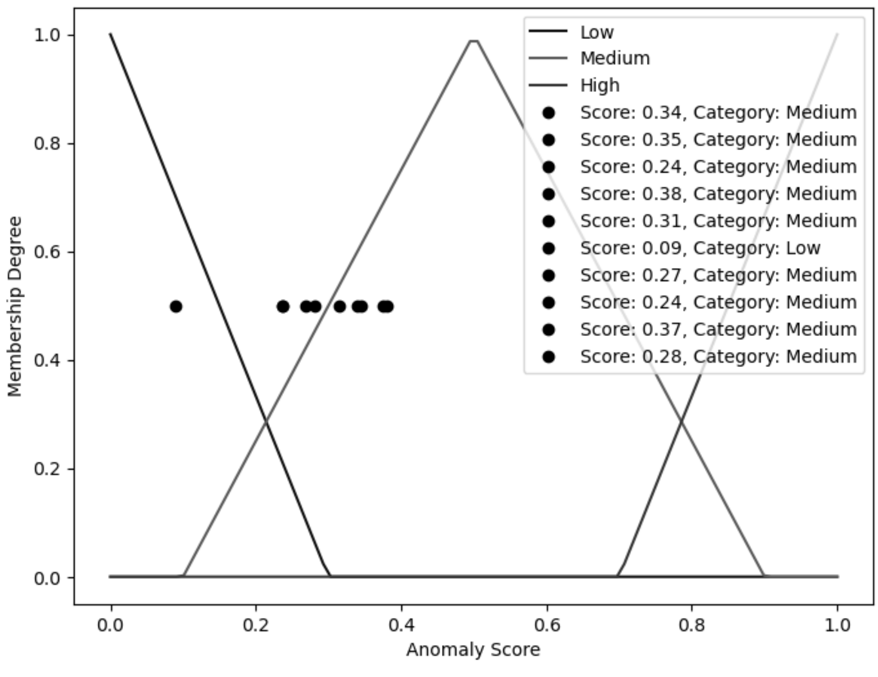


Рис. 9. Нечітка оцінка

Розглянута система виявлення аномалій є перспективною, але, тим не менш, має певні недоліки та потребує вдосконалення. ГЗМ значною мірою залежать від різноманітності аномалій у навчальних даних та збалансованості даних; якщо різноманітності та збалансованості бракує, це може негативно відобразитись на здатності системи до генералізації. Крім того, є проблеми масштабованості даних, чутливості до шуму, а також обмеження у виявленні нових аномалій [17]. Усунення цих обмежень вимагатиме подальших досліджень, удосконалення моделі та індивідуальної адаптації для конкретних застосувань.

Висновки

ГЗМ продемонстрували свою здатність до високоякісної обробки і генерації графічних даних. Крім того, їх архітектурна особливість дозволяє виходити за рамки лише генеративних функцій моделі та застосовувати ГЗМ у незвичних задачах, як, наприклад, виявлення аномалій.

Також цей тип нейронних мереж має потенціал у широкому спектрі застосування в природничих науках і технологічних розробках.

Попри свою ефективність, ГЗМ все ще мають недоліки, найсуттєвішим з яких є трудомісткість процесу навчання і процес налаштування гіперпараметрів. Саме цей аспект ГЗМ представляє найбільший інтерес для досліджень, оскільки розв'язання цієї проблеми потенційно може сильно підвищити можливості ГЗМ і генеративного ШІ зокрема.

Автори продовжать дослідження та експерименти з ГЗМ для подальшого вдосконалення їх навчання і підвищення ефективності їх генеративного компонента.

Література

1. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, J. Bengio, “Generative Adversarial Networks,” *Proceedings of the International Conference on Neural Information Processing Systems* (NIPS) 2014, pp. 2672—2680.
2. N. Aldausari, A. Sowmya, N. Marcus, and G. Mohammadi, *Video Generative Adversarial Networks: A Review*, 2022, [Online]. Available at: https://doi.org/10.1145/3487891
3. O. S. Striuk, Y. P. Kondratenko, “Generative Adversarial Neural Networks and Deep Learning: Successful Cases and Advanced Approaches,” *International Journal of Computing*, vol. 20, issue 3, pp. 339-349, 2021.
4. O. S. Striuk, Y. P. Kondratenko, Generative Adversarial Networks in Cybersecurity: Analysis and Response, in: Y.Kondratenko, V.Kreinovich, W.Pedrycz, A.Chilrii, A.M.Gil-Lafuente (Eds.), Artificial Intelligence in Control and Decision-making Systems: Dedicated to Prof. Janusz Kacprzyk. Studies in Computational Intelligence, vol. 1087, Springer, Cham, 2023, pp. 373-388.
5. M. Arjovsky, L. Bottou, Towards Principled Methods for Training Generative Adversarial Networks, 2017, [Online]. Available at: https://arxiv.org/abs/1701.04862
6. O. S. Striuk, Y. P. Kondratenko, “Optimization Strategy for Generative Adversarial Networks Design,” International Journal of Computing, vol. 22, issue 3, pp. 292-301, 2023.
7. R. Ayari, Generative Adversarial Networks, 2020, [Online]. Available at: https://bit.ly/3Uk4GBw
8. A. Borji, Pros and Cons of GAN Evaluation Measures, 2018, [Online]. Available at: https://arxiv.org/abs/1802.03446
9. J. Brownlee, A Gentle Introduction to Transfer Learning for Deep Learning, 2017, [Online]. Available at: https://bit.ly/3GTmdeC
10. O. Striuk, Y. Kondratenko, I. Sidenko, A. Vorobyova, “Generative Adversarial Neural Network for Creating Photorealistic Images,” Proceedings of 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory, Kyiv, Ukraine, November 27, 2020, pp. 368-371.
11. Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, “Gradient-based learning applied to document recognition,” in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
12. O. Striuk and Y. Kondratenko, “Adaptive Deep Convolutional GAN for Fingerprint Sample Synthesis,” 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT), Lviv, Ukraine, 2021, pp. 193-196, doi: 10.1109/AICT52120.2021.9628978.
13. A. Bécue and C. Champod, "Interpol review of fingermarks and other body impressions (2019 – 2022)," Forensic Science International: Synergy, vol. 6, p. 100304, 2023.
14. O. Striuk and Y. Kondratenko, “Implementation of Generative Adversarial Networks in Mobile Applications for Image Data Enhancement,” Journal of Mobile Multimedia, vol. 19, no. 03, pp. 823–838, 2023. doi: 10.13052/jmm1550-4646.1938.
15. Y. I. Shehu, A. Ruiz-Garcia, V. Palade, A. James, “Sokoto Coventry Fingerprint Dataset,” arXiv:1807.10609 [cs.CV], 2018, pp. 1—3.
16. BIRDS 400 Dataset, [Online]. Available at: https://www.kaggle.com/datasets/gpiosenka/100-bird-species.
17. F. Di Mattia et al., A Survey on GANs for Anomaly Detection, 2021, [Online]. Available at: https://arxiv.org/abs/1906.11632

References

1. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, J. (2014). “Generative Adversarial Networks,” *Proceedings of the International Conference on Neural Information Processing Systems* (NIPS), 2672-2680.
2. Aldausari, N., Sowmya, A., Marcus, N. and Mohammadi, G. (2022). *Video Generative Adversarial Networks: A Review* [Online]. Available at: https://doi.org/10.1145/3487891
3. Striuk, O.S., Kondratenko, Y.P. (2021). “Generative Adversarial Neural Networks and Deep Learning: Successful Cases and Advanced Approaches,” *International Journal of Computing*, vol. 20, issue 3, 339-349.
4. Striuk, O.S., Kondratenko, Y.P. (2023). Generative Adversarial Networks in Cybersecurity: Analysis and Response, in: Y. Kondratenko, V. Kreinovich, W. Pedrycz, A. Chilrii, A. M. Gil-Lafuente (Eds.), Artificial Intelligence in Control and Decision-making Systems: Dedicated to Prof. Janusz Kacprzyk. Studies in Computational Intelligence, vol. 1087, Springer, Cham, 373-388.

5. Arjovsky, M., Bottou, L. (2017). Towards Principled Methods for Training Generative Adversarial Networks [Online]. Available at: https://arxiv.org/abs/1701.04862

6. Striuk, O.S., Kondratenko, Y.P. (2023). “Optimization Strategy for Generative Adversarial Networks Design,” International Journal of Computing, vol. 22, issue 3, 292-301.

7. Ayari, R. (2020). Generative Adversarial Networks [Online].
Available at: https://bit.ly/3Uk4GBw

8. Borji, A. (2018). Pros and Cons of GAN Evaluation Measures [Online]. Available at: https://arxiv.org/abs/1802.03446

9. Brownlee, J. (2017). A Gentle Introduction to Transfer Learning for Deep Learning [Online]. Available at: https://bit.ly/3GTmdeC

10. Striuk, O., Kondratenko, Y., Sidenko, I., Vorobyova, A. (2020, November 27). “Generative Adversarial Neural Network for Creating Photorealistic Images,” Proceedings of 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory, Kyiv, Ukraine, 368-371.

11. Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). “Gradient-based learning applied to document recognition,” in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, doi: 10.1109/5.726791.

12. Striuk, O. and Kondratenko, Y. (2021). “Adaptive Deep Convolutional GAN for Fingerprint Sample Synthesis,” 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT), Lviv, Ukraine, 193-196.
doi: 10.1109/AICT52120.2021.9628978.

13. Bécue A. and Champod, C. (2023). "Interpol review of fingermarks and other body impressions (2019–2022)," Forensic Science International: Synergy, vol. 6, 100304.

14. Striuk, O. And Kondratenko, Y. (2023). “Implementation of Generative Adversarial Networks in Mobile Applications for Image Data Enhancement,” Journal of Mobile Multimedia, vol. 19, 03, 823–838, doi: 10.13052/jmm1550-4646.1938.

15. Shehu, Y.I., Ruiz-Garcia, A., Palade, V., James, A. (2018). “Sokoto Coventry Fingerprint Dataset,” arXiv:1807.10609 [cs.CV], 1-3.

16. BIRDS 400 Dataset, [Online]. Available at: https://www.kaggle.com/datasets/gpiosenka/100-bird-species.

17. Di Mattia F. et al. (2021). A Survey on GANs for Anomaly Detection [Online]. Available at: https://arxiv.org/abs/1906.11632

The article has been sent to the editors 15.10.23.
After processing 25.10.23.
Submitted for printing 30.11.23.